

中华医学会系列杂志及作者支持 原始数据共享政策的调查分析

季媛媛¹ 陈立敏¹ 刘冰² 包雅琳¹ 齐文安^{3†}

1) 中华医学会杂志社《中华医学杂志》(英文版)编辑部; 2) 中华医学会杂志社; 3) 中华医学会杂志社《英国医学杂志》(中文版)编辑部; 100710, 北京

摘要 原始数据共享能提高学术诚信、减少科研浪费。期刊有义务帮助促进数据共享。本文通过对中华医学会系列杂志在2016年中国科技论文统计源期刊目录中的111种期刊,以及《中华医学杂志》(英文版)2016年发表文章的438名作者进行调查,了解中华医学会系列杂志及作者对原始数据共享政策的支持情况,并与BMJ出版集团出版的62种期刊(BMJ Journals)相比较,探索中国医学期刊推动数据共享政策的途径,提出中国医学科技期刊可从发表相关述评类或观点类文章、制定相应的编辑政策、开展丰富的互动活动等方面来推动原始数据共享。

关键词 数据共享; 医学期刊; 编辑政策

Endorsement of data sharing by journals hosted by Chinese Medical Association and authors in China//JI Yuanyuan, CHEN Limin, LIU Bing, BAO Yalin, QI Wen'an

Abstract Sharing data will increase confidence and trust from researches and avoid unwarranted repetition. Journals have obligation to help foster data sharing. This study investigated the websites of 111 high-impact journals hosted by the Chinese Medical Association and distributed an electronic questionnaire to 438 corresponding authors of *Chinese Medical Journal* who published manuscripts in 2016, to assess the endorsement of data sharing by authors and medical journals in China, compared with 62 journals published by BMJ Journals. We also suggest medical journals in China to publish editorials or viewpoints on data sharing, to change their requirements for publishing manuscripts, and to develop interactive activities with authors to help foster data sharing.

Keywords data sharing; medical journal; editorial policy

First-author's address Chinese Medical Journal, Chinese Medical Association, 42 Dongsixidajie, 100710, Beijing, China
DOI:10.16811/j.cnki.1001-4314.2018.01.015

医学科学数据是开展医学研究工作的重要知识来源。医学科学数据共享在加强学术研究、医学实践和临床试验系统的完整性方面具有巨大的潜力,能够提高学术诚信并减少不必要的重复性研究、促进不同研究团队和不同学科之间的合作,有利于医学知识的传播与利用。近年来,在开放获取理念的指导下,国际组织及各国政府积极推动其资助的科研产出的开放获取,特别是在医学领域,以美国国立卫生研究院为代表

的医学研究机构也纷纷制定了科学数据共享的开放获取政策^[1]。

WHO 临床试验注册平台于2015年8月发布了关于支持临床试验数据共享的声明。国际医学期刊编辑委员会(International Committee of Medical Journal Editors, ICMJE)于2016年1月20日发布了关于临床试验数据共享的建议,要求研究者在临床试验投稿时提供其原始数据在公共数据库的信息,供编辑、审稿专家和公众共享,并将其作为在成员期刊中考虑发表的必要条件^[2]。共享临床试验原始数据已经成为临床试验透明化的重要组成部分。

全球4大综合性医学期刊之一的《英国医学杂志》(The BMJ)近期报道:中国国家食品药品监督管理局(CFDA)于2016年9月9日就其接收的1622种注册新药相关数据进行调查,结果显示其中1308种药物临床试验数据涉嫌伪造、不完整或不充分等问题,中国超过一半临床试验数据涉嫌造假^[3]。中国医学临床试验在世界期刊中的诚信度已受到质疑,中国医学科技期刊有责任和义务帮助促进中国医学研究数据共享、增强试验数据透明、提高中国医学科研诚信、杜绝科研造假。

本研究旨在通过调查中华医学会系列杂志及作者人群对原始数据共享政策的支持情况,了解中国作者对数据共享资源数据库的运用及对原始数据共享政策的顾虑,并与BMJ出版集团出版的期刊(BMJ Journals)相比较,探索中国医学期刊推动数据共享政策的途径。

1 资料和方法

1.1 研究的主要内容 审阅中华医学会系列杂志在2016年中国科技论文统计源期刊目录中的111种期刊,以及英国医学会旗下BMJ出版集团出版的62种期刊的官方网站,提取信息。对《中华医学杂志》(英文版)(*Chinese Medical Journal*)2016年发表文章的438名作者进行问卷调查。

1.2 研究方法 本次调查开展时间为2017年2月1—15日,分为以下2个部分:1)审阅111种中华医学

† 通信作者

会系列杂志及 BMJ 出版集团出版的 62 种期刊官网中的“稿约”及“编辑部公告”,检查期刊是否介绍了数据共享相关概念信息,以及是否要求作者对原始数据进行共享。2)对《中华医学杂志》(英文版)的 438 名通信作者发送邮件进行问卷调查。问卷主要包括 6 方面的问题:①作者是否在国外 SCI 期刊中发表过文章?②在此次调查之前,作者是否知晓原始数据共享政策?③作者是否支持原始数据共享政策?④作者是否有过上传/下载原始数据行为?⑤作者对现行的数据共享资源数据库是否有所了解?⑥如对原始数据共享存在顾虑,原因是什么?

1.3 数据统计 计算作者人群对临床试验原始数据共享政策的知晓率与支持率、数据共享行为发生率,以及对现行数据共享资源数据库的了解情况。

通过问卷调查的“问题 1”将作者分为“有国外 SCI 期刊发表经验者”与“仅在国内期刊发表过文章者”2 组,通过卡方检验比较 2 组作者对原始数据共享的知晓率、支持率及数据共享行为方面的差别;通过 Kruskal-Wallis 检验比较 2 组作者对数据共享资源数据库运用方面的差别, $P < 0.05$ 视为差异有统计学意义。

2 结果与分析

2.1 中华医学会系列杂志与 BMJ 出版集团期刊对原始数据共享政策的支持情况

2.1.1 111 种中华医学会系列杂志对原始数据共享政策的支持情况 调查的 111 种中华医学会系列杂志中仅有 2 种(1.8%, 2/111)期刊(《中华外科杂志》与《中华神经外科杂志》(英文版)在编辑政策公告中建议作者共享原始数据并提供了期刊接受何种共享方式等信息。

《中华外科杂志》建议作者将图片等原始资料放在稿件末尾,标明“以下为本研究的原始资料”,并简要描述各图片的具体内容;而对数据较多的文章,建议作者将原始数据上传云盘并附云盘链接地址。《中华神经外科杂志》(英文版)按照其出版商 BioMed Central 的要求,鼓励作者将其原始数据存放在公共资源数据库中,或者以补充支持材料的形式提交;要求作者在文章中注明“数据和材料的可用性”,详细说明原始数据的存储地址和获得方式,如作者不希望共享原始数据也必须说明并给出合理的原因。

2.1.2 62 种 BMJ 出版集团期刊对原始数据共享政策的支持情况 BMJ 出版集团出版的期刊中有 54 种(87.1%, 54/62)在稿约或编辑政策公告中建议作者共享原始数据。未建议作者共享原始数据的 8 种期刊

为单体病例报告、不接受外来投稿、不接受研究论文投稿或兽医类杂志。

这 54 种支持原始数据共享政策的期刊中,《The BMJ》作为 ICMJE 成员刊之一,对原始数据共享政策要求最为严格。《The BMJ》要求对药物或设备的临床试验必须共享原始数据,其他研究类型鼓励作者共享原始数据。要求作者在每篇文章末尾按照建议的格式选项陈述“数据共享声明”,内容包括:数据共享的内容、资源数据库网址及 doi 号、数据共享的形式,以及数据共享知情同意。如无原始数据用于共享,也应在文末注明(但此选项对临床试验不适用)。

其他 53 种支持原始数据共享的期刊均表示鼓励原创性研究文章的作者在投稿时递交一份数据共享声明。声明须阐明如有额外未公开发表的数据,哪些部分可以公开,向谁公开,如何获取。BMJ 出版集团推荐作者在 Dryad(<http://datadryad.org/>)上进行数据共享,但其也注明 Dryad 是完全无限制性共享数据,研究者应根据实际情况决定是否采用这个共享数据库。

2.2 中国作者人群对原始数据共享政策的支持情况

2.2.1 调查问卷回复情况 截止到 2017 年 3 月 15 日,共收到 247/438(56.4%)位作者提交的调查问卷,其中 132 名作者有国外 SCI 期刊发表经验,而 115 名作者仅在国内期刊发表过文章。

2.2.2 中国作者对原始数据共享政策的知晓及运用情况 仅 1/3 的作者(88/247, 35.6%)反馈在此次调查前知晓原始数据共享的概念。大多数作者(215/247, 87.0%)都表示支持原始数据共享;只有 11.7%的作者(29/247, 11.7%)曾有过共享原始数据行为。其中,仅有极少数作者(7/247, 2.8%)能够熟练运用数据共享资源数据库,41.7%的作者对数据共享资源数据库仅有了解(103/247, 41.7%),而超过半数作者对此完全不知晓(137/247, 55.5%)。

2.2.3 2 组作者对原始数据共享政策的知晓及运用情况的差别 与有国外 SCI 期刊发表经验的作者比较而言,仅在国内期刊发表过文章的作者显示出更低的数据共享知晓率及数据共享行为发生率,且差异有统计学意义;但 2 组作者对原始数据共享的支持率相似,且对数据共享资源数据库的掌握程度差异无统计学意义(表 1)。

2.2.4 中国作者对原始数据共享政策存在的顾虑 作者对原始数据共享存在的顾虑选择最多的是“可能导致科研成果被窃取”,占 29.2%,24.3%作者顾虑“研究者滥用共享数据”,23.5%作者担心“增加侵犯临床试验参与者的隐私风险”,15.5%的作者表示“不

了解数据共享资源数据库”,7.5%的作者存在除以上列举的其他顾虑。

表1 有国外SCI期刊发表经验与仅在国内期刊发表过文章的作者之间对数据共享知晓及运用情况的比较

问卷项目	有国外SCI期刊发表经验作者 (n=132)	仅国内期刊发表文章作者 (n=115)	χ^2 值	P 值
知晓数据共享	57(43.2%)	31(27.0%)	7.055	0.008
支持数据共享	110(83.3%)	105(91.3%)	3.463	0.063
有数据共享行为	21(15.9%)	8(7.0%)	4.753	0.029
数据共享资源数据库运用			1.441	0.230
熟练运用	3(2.3%)	4(3.5%)		
仅有了解	61(46.2%)	42(36.5%)		
未听说过	68(51.5%)	69(60.0%)		

2.3 数据分析 从调查数据中可以看到,目前我国医学期刊对原始数据共享政策的介绍和支持远低于国际医学期刊。而从我国作者角度来看,绝大多数作者都支持原始数据共享政策以增强科研透明度、杜绝科研造假,但国内作者对原始数据共享政策的认知度以及普及性偏低,且普遍对数据共享资源数据库不甚了解。有国外SCI期刊发表经验的作者比仅在国内期刊发表过文章的作者对原始数据共享有明显更高的认知度和执行率。这提示我们,国内医学期刊应做更多工作来推动临床试验原始数据共享,保持与国际期刊政策接轨。

3 讨论

3.1 医学研究原始数据共享的意义 无论从作者个人层面、编辑和审稿人审查稿件层面、基金资助单位方面,还是对于科研及发表伦理、公众健康以及大数据需求价值方面,医学数据共享都有重要积极的意义。

数据共享可提高科学数据的利用率 and 价值,促进不同研究团队和不同学科之间的合作,更好地佐证自己的研究,也提高了作者个人学术成果的影响力;作者投稿时在公共数据库提交原始数据,可供医学期刊编辑和审稿专家审核稿件时核查,保证了研究结果的真实性,减少学术不端对期刊出版事业的影响;项目管理机构或基金资助方可根据需要随时了解医学研究的执行情况和进度,满足目前诸多研究资助者鼓励甚至强制数据共享的要求^[4];医学原始数据共享能够减少科研工作的重复性及不必要的科研浪费,推动临床试验透明化,让公众了解临床试验执行细节和过程,保证临床试验资料的真实性,增加临床试验的公信度,提高科研诚信;数据共享也体现了医学研究者的社会责任感,对承担安全风险、为全人类健康事业做出卓越奉献的参试者的尊重,便于公众了解医学临床试验的所有细

节和结果。此外,在大数据爆发性增长的今天,基于互联网的全国性或全球性协作临床试验已经在国际上广泛开展,医学原始数据共享有利于数据的长期保存和使用,可实现医疗卫生资源利用最大化,为高质量研究提供证据,推进医学事业的全球化进程,是未来临床试验大数据时代的必然趋势。

3.2 中国医学科技期刊应从哪些方面推动数据共享

数据共享是一个共同责任,中国医学科技期刊亦有责任和义务帮助促进数据共享,可从以下方面来具体实施。

3.2.1 发表相关述评类或观点类文章,提倡原始数据共享 《The BMJ》是最早要求数据共享的期刊之一,于2015年5月7日刊发述评要求所有临床试验都须按要求共享原始数据^[5],鼓励作者投稿时标注数据共享声明,包括本研究未出版的其他数据、谁能获取这些数据以及如何获得数据等内容。《新英格兰医学期刊》(NEJM)于2016年5月起刊发多篇述评及观点,阐明遵循ICMJE关于临床试验数据共享的建议,呼吁作者共享原始数据^[6-7]。这些评论类文章可提高读者人群对原始数据共享的认知度。

2017年6月,ICMJE发表了一项对临床试验数据共享声明的统一要求^[8],将以下内容作为在其成员刊发表临床试验报告的考虑条款:1)对于2018年7月1日及以后提交到ICMJE成员刊的临床试验报告,必须包含数据共享声明;2)对于2019年1月1日及以后开始入组受试者的临床试验,必须在临床试验注册平台上提交数据共享计划。如果数据共享计划有变化或更改,应在注册平台上进行更新并在提交论文时加以说明。

ICMJE数据共享声明包含以下5个方面的内容:是否将共享研究对象的去标识个体数据(包括变量清单(data dictionaries));会共享哪些数据;是否可以获得其他与研究相关的文档(研究方案、统计分析计划等);何时可获取数据以及可开放获取多久;获取共享试验数据的要求(包括谁能获取数据、用于何种研究目的,以及如何获取)。

尽管这些初步要求尚未强制进行数据共享,但ICMJE提醒作者,编辑会参考作者提供的数据共享声明来做出稿件最终是否刊出的决定。一些ICMJE成员期刊已经接受或采取更为严格的数据共享要求。

3.2.2 制定相应的编辑政策,促进原始数据共享

1)在医学期刊“投稿须知”或栏目公告中介绍数据共享的相关知识,提供数据存储资源库链接,帮助作者根据自身研究类型选择适合的数据库,这些措施可提高作者人群对原始数据共享以及存储数据库运用的认知度。

2)参考国际期刊做法,改变稿件发表要求,鼓励

原始数据共享。《Annals of Internal Medicine》《The BMJ》《PLoS Medicine》等杂志在编辑政策中表明,出版前一旦作者拒绝分享数据、相关元数据和方法,编辑有理由拒稿;出版后一旦期刊发现无法正常获取原始数据,期刊可以选择要求作者提供更多数据信息、发表关注声明、通知其研究基金资助机构,或在某些情况下撤销出版。在极少数确实不可能遵守数据共享要求的情况下,编辑可以考虑作者的豁免请求,如果接受数据共享豁免,必须在出版物中解释原因^[9]。

3.2.3 开展丰富的互动活动,推动原始数据共享

1)编辑部可在期刊官网中开辟专栏解答作者对数据共享的疑惑,促进作者及读者人群对数据共享的深入熟悉;2)编辑部尝试将推广原始数据共享及培训作为工作内容之一,加强教育与培训,呼吁作者遵循数据共享政策;3)编辑部积极探索不同类型数据共享的标准、形式规范以及质控方式。科技部支持的国家人口与健康科学数据共享平台的专题数据服务平台是中英双语对国内外开放的科学数据网络共享平台。2017年1月,该平台将设置专门的肿瘤专题平台版块发布《中华外科杂志》的文章数据(发布内容和格式包括文章名称、摘要、关键词、支持数据)。2017年以前已发表的文章,根据自愿原则,文章通信作者可以提交去隐私化的支持文章结果和结论的原始数据。肿瘤专题平台将对数据进行审核,符合要求的数据在平台发布。2017年1月1日起,在《中华外科杂志》投稿的稿件,采取自愿原则,提倡作者以数据自身格式为基础,在投稿的同时提交去隐私化支撑数据。肿瘤专题平台组织专家,初步审核数据的规范性、完整性,以及对所投稿件结果和结论的相关性等初步意见,该意见不作为稿件是否录用的条件。

3.2.4 呼吁搭建完善中国医学数据共享平台,为原始数据共享奠基 我国医学数据共享网站可见性略低于其他学科数据共享平台,原因主要是资源量少或是存在部分资源链接无法访问的情况。我国医学期刊可呼吁国家搭建完善中国自己的医学数据共享平台,关注网络影响力的提高,加强元数据建设,注重网络性能提升,重视数据资源更新维护^[10]。2017年1月4日,国家人口与健康科学数据共享平台科技资源发布会在京举行,会议主题为“医学科学数据共享,推进健康中国建设”。会上发布了2016年度国家人口与健康科学数据共享平台科学数据资源,共发布包括生物医学、基础医学、临床、公共卫生、中医药学、药学、人口与生殖健康7大类237个数据集,数据量达到49.1TB,平台为政府卫生决策、医疗卫生事业发展、科技创新、应对突发公共卫生事件和百姓健康等提供了巨大的科技支撑作用。

3.2.5 关注医学数据共享监管法规,为原始数据共享护航 我国医学期刊可学习关注国际医学期刊数据共享的监管法规和实践经验,合理管理、使用作者上传的原始数据,保护数据安全、隐私,避免数据滥用、误用。公共科学图书馆·综合(Public Library of Science,简称PLOS ONE)是较早提出科学数据共享政策的学术期刊之一,其科学数据共享政策影响较广、实践较早且数据共享较为规范。针对数据存储及附件数据因涉及道德、法律等问题无法公开数据的情况,《PLOS ONE》提供了“依申请公开”及“正确引用第三方数据”这2种可选的方法进行评估。在第1种情况下,数据可用性声明中必须详细说明“数据基于请求可用”,说明限制数据公开存储的原因,并明确研究者应将请求提交给如数据访问委员会等团体;第2种情况适用于论文中的主数据集并非由作者自己生成,《PLOS ONE》要求有需求的研究人员可从指定的原始来源中独立地获取第三方数据。针对人类受试者研究的临床数据、临床试验等病人隐私保护,《PLOS ONE》提出了明确的要求:作者必须避免公开能识别病人身份的信息,除非有严格的提交要求^[11]。

医学原始数据共享政策须保护作者的原创成果和受试者的隐私,推动医学数据共享不仅需要期刊来推动作者加强共享意识、呼吁完善坚实的技术基础,更需要期刊主管部门制定数据共享管理政策法规体系来约束与规范共享行为。期刊主管部门应制定国内数据共享指南,健全共享法规体系,明确数据共享原则,界定数据共享范围、程度,明晰大数据产权和知识产权的界定,建立数据共享利益协调机制^[12-13]。数据共享法规体系应强调对所存储数据的开放利用及在利用过程中对数据生产者的权利保护,包括数据所有权管理、开放共享、许可协议、引用规范等。完善的原始数据共享法规可解决国内作者对共享原始数据存在的顾虑,为推动数据共享保驾护航。

虽然短时间内要求国内医学科技期刊发表文章均遵循ICMJE的建议、实现临床试验原始数据共享尚不契合国内现况实际,但我国医学期刊可以制定优先发表政策等激励措施,优先发表接受原始数据共享的研究论文,提高医学研究者共享原始数据的热情。期刊人应适应时代潮流,紧追国际趋势,保持出版理念领先。

4 结束语

医学科技期刊实施原始数据共享将提高临床试验结果的可信度、促进新医学假说的发展和检验。共享临床试验原始数据将避免不必要的重复研究、使每个临床试验的数据得到更高效的利用。临床试验数据共